WP 37
26 April 2010

UNITED NATIONS STATISTICAL COMMISSION and ECONOMIC COMMISSION FOR EUROPE

STATISTICAL OFFICE OF THE EUROPEAN UNION (EUROSTAT)

CONFERENCE OF EUROPEAN STATISTICIANS

Joint Eurostat/UNECE Work Session on Demographic Projections
(28-30 April 2010, Lisbon, Portugal)

Item 12 – Stochastic techniques for demographic projections

Bayesian Population Forecasts for England and Wales[1]

Guy J. Abel[2], Jakub Bijak, Jonathan J. Forster, James Raymer and Peter W.F. Smith,
University of Southampton, United Kingdom

ABSTRACT

The Bayesian approach has a number of attractive properties for forecasting uncertainty which have yet to be fully explored in the study of future population change. In this paper, we apply some simple Bayesian time series models to obtain future population estimates with uncertainty for England and Wales. Uncertainty in model choice is incorporated through Bayesian model averaging techniques. The resulting predictive distributions from Bayesian forecasting models have two main advantages over those obtained using more traditional stochastic models. First, uncertainties in the data, the model parameters and model choice are explicitly represented using probability distributions. As a result, more realistic probabilistic population forecasts are obtained. Second, Bayesian models formally allow the incorporation of expert opinion, including uncertainty, into the forecast. We conclude by discussing our results in relation to classical time series methods and existing cohort component estimates.

1. INTRODUCTION

This paper explores the use of Bayesian methods for population forecasting. The main rationale for this paper is rooted in two philosophical debates of vital importance for demographers. The first one concerns the need for incorporating uncertainty in population forecasts, advocated by many authors since the 1980s (e.g. Alho and Spencer, 1985; Keyfitz, 1991; Lee 1998), as opposed to deterministic scenario projections with a (near) zero probability of realisation. The second debate relates to the issue of simplicity versus complexity in demographic forecasting (e.g. McNown et al., 1995).

We believe there are three important considerations that could improve our understanding about population forecasts and the various sources of uncertainty about them. First, the Bayesian approach offers an explicit, coherent and transparent mechanism to include uncertainty in the data, parameters of the model and the model itself, by using probability distributions. Second, it allows the inclusion of expert judgements, including uncertainty, into the model framework. Third, the predictive distributions simply follow from the probabilistic model applied. As a result, probabilistic population forecasts, with more reliable and coherent estimates of predictive distributions, can be obtained. Together, these have the potential to improve the measurement of uncertainty in forecasts, and thus improve our potential for planning and understanding population change.

To investigate the potential of a Bayesian approach, we apply time series models to prepare forecasts up to 2033 based on a simple data set of population totals. Uncertainty in model choice is incorporated through Bayesian model averaging techniques. The results and their levels of uncertainty are compared with the traditional frequentist methods of time series analysis and model selection. Further contrasts are also made against the Office of National Statistics (ONS) national population projections which rely on deterministic scenarios and cohort component methodologies.

---

[1] This research is funded by ESRC Grant number RES-625-28-0001. The Centre for Population Change is a joint initiative between the University of Southampton and a consortium of Scottish Universities in partnership with the Office for National Statistics (ONS) and the General Register Office for Scotland (GROS). The findings, interpretations, and conclusions expressed in this paper are entirely those of the author(s) and should not be attributed in any manner to ONS or GROS.

[2] Contact author: ESRC Centre for Population Change, University of Southampton, Southampton SO17 1BJ, United Kingdom; email G.J.Abel@soton.ac.uk.

The paper is structured as follows. In the next section, we introduce our data, a time series of data for England and Wales from 1841 to 2007. In Section 3 we introduce the notation, and describe the models used in this study. These include autoregression models for time series, the Bayesian inference used for parameter calculation and model averaging, which is applied to account for model uncertainty and to provide a more robust set of estimates. In Section 4, we apply the models of the previous section to forecast the populations in England and Wales from 2008 to 2033. Section 5 compares these empirical forecasts with the outcome from two other methods, namely traditional frequentist methods of time series analysis and the official, deterministic population projections for England and Wales prepared by the ONS. Finally, we end the paper with a summary and discussion, which include suggestions for extending the proposed approach to more complex population models. These include, for example, models to account for heterogeneity found in the historical data and multivariate models for time series of birth, death and migration rates.

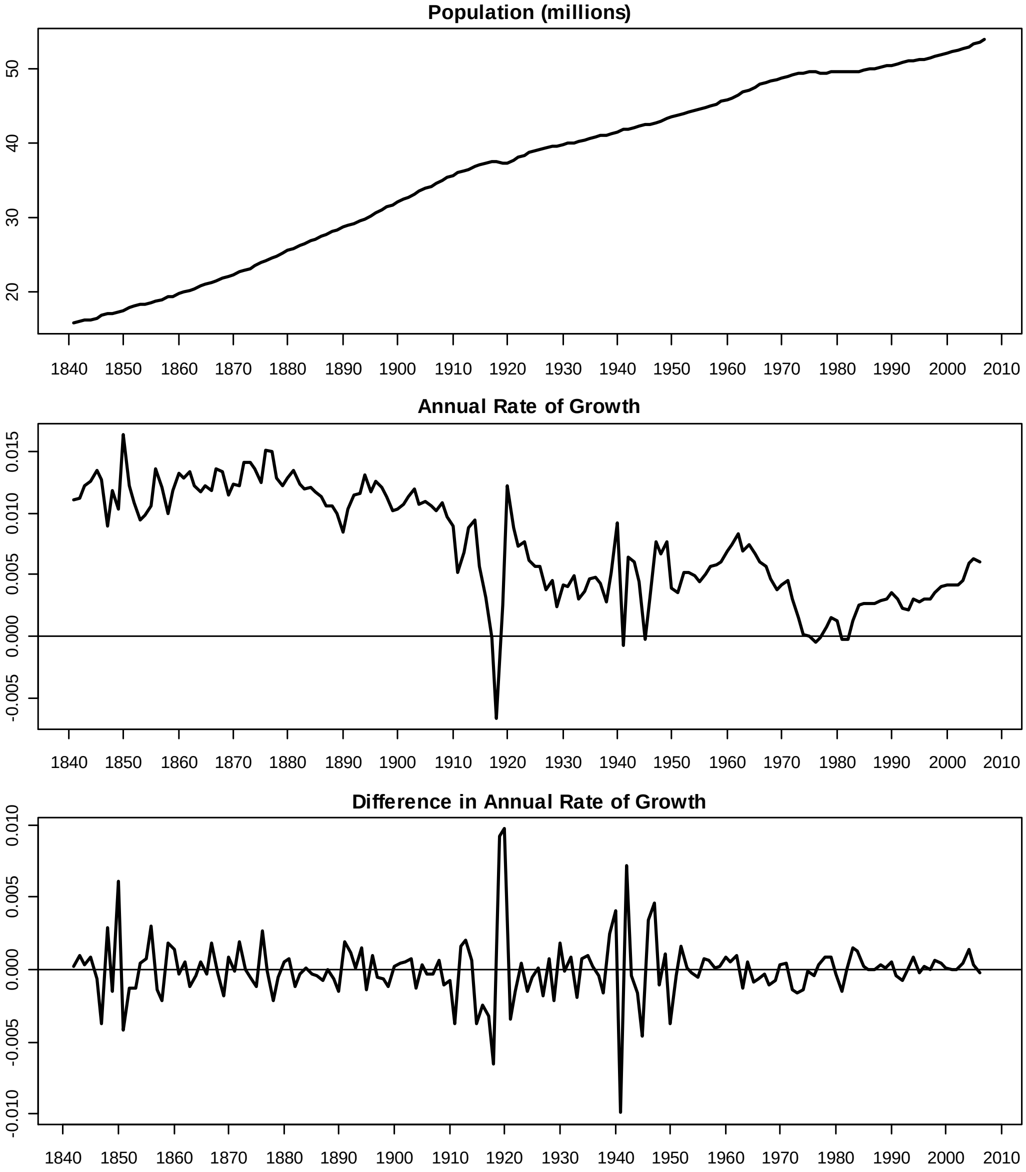
2. DATA

To illustrate a Bayesian approach to modelling population data, we focus on the simplest case, i.e., a time series of population estimates obtained from the Human Mortality Database (2009) for England and Wales from 1841 to 2007. The mid-year population totals (including military personnel) are presented in the top panel of Figure 1.

The England and Wales population totals exhibited a steady increase over time, rising from 15.8 million in 1841 to 53.9 million in 2007. A slight decrease is noticeable from the effects of the First World War and the 1918 influenza pandemic. Also noticeable is the slower increase in the population during the late 1970s and early 1980s, a result of net emigration and a slow rate of natural population increase.

The features of population change are more evident when the rates of growth, plotted in the second panel of Figure 1, are considered. In addition, a number of other characteristics become apparent. During the first third of the series the population grew at higher rates than any of the remaining time period. This was predominantly due to the declining mortality occurring substantially before a decline in fertility, which remained at pre-industrial levels for much of this period. Between the two wars the rate of growth remained low in comparison with the later half of the 19th century and early 20th century. This was partly driven by low fertility levels being reached through the economic depression and a change in sociological factors. After the Second World War population growth rates increased as a result of erratic fluctuations in fertility. This occurred initially through a short lived baby boom associated with demobilisation followed by a more substantial increase in fertility during the 1950s and early 1960s. In the late 1970s and early 1980s, the previously mentioned low (and occasionally negative) levels of growth took place followed by a rise in recent decades driven by net immigration and increased fertility levels. For a detailed discussion of England's population history, see Coleman and Salt (1992) and Hinde (2003).

Figure 1: England and Wales Population Data, 1841-2007



3. MODELS

Let p_t be the population size at t and $t = 1,\ldots,T$ represent an uninterrupted series of time points at which p_t is observed. The problem of population forecasting is to obtain estimates of p_t for one of more values of $t > T$. Furthermore, the utility of the forecasts is enhanced when associated measures of uncertainty are provided.

In order to model p_t, we first derive the time series of population increments r_t where

$$p_{t+1} = (1 + r_t)\,p_t\,. \tag{1}$$

However, experience suggests (Chatfield p26, 2004) that if we are to use models which assume stationarity, it is more appropriate to model changes in r_t:

$$y_t = r_t - r_{t-1}\,. \tag{2}$$

For our data, time series plots of r_t and y_t are presented in the second and third panels, respectively, of Figure 1. In the following subsections we introduce autoregression (AR) models for y_t.

3.1 Autoregression Models

AR models have been used in the demographic context to forecast population (e.g. Saboia (1974), Ahlburg (1987), Pflaumer (1992), Alho and Spencer (2005)). An AR model of order *p*, denoted AR(*p*), is defined as

$$y_t = \sum_{k=1}^{p} \phi_k y_{t-k} + z_t \, , \tag{3}$$

where z_t are independent observations from a probability distribution with zero mean and constant variance, σ^2. A slightly more flexible model, which we shall use, also allows for a non-zero mean for y_t and is defined as

$$y_t = \mu + \sum_{k=1}^{p} \phi_k (y_{t-k} - \mu) + z_t \, . \tag{4}$$

This model for y_t implies a mean increase in r_t of μ each year. For a fully-specified probability model, we need to assume a distribution for z_t. Typically, a normal distribution is assumed.

3.2 Bayesian Inference

In Bayesian inference, uncertainty about the (multivariate) parameter θ of a statistical model is described by its posterior probability distribution given observed data $y_{\{T\}} = (y_1, \ldots, y_T)$. The probability density function of y_t is obtained by using Bayes Theorem:

$$f(\theta \mid y_{\{T\}}) = \frac{f(y_{\{T\}} \mid \theta) f(\theta)}{f(y_{\{T\}})} \, , \tag{5}$$

where $f(y_{\{T\}} \mid \theta)$ is the likelihood function and is defined by the model, $f(\theta)$ is the prior distribution for θ and $f(y_{\{T\}})$ is a normalising constant. The prior distribution $f(\theta)$ specifies the uncertainty about θ prior to observing any data.

Forecasting or prediction is particularly natural in a Bayesian framework. Uncertainty about the next *K* future values of y_t (for $t = T+1, \ldots, T+K$) is described by the joint predictive probability distribution

$$f(y_{T+1}, \ldots, y_{T+K} \mid y_{\{T\}}) = \int f(\theta \mid y_{\{T\}}) \prod_{k=1}^{K} f(y_{T+K} \mid y_{T+K-1}, \theta) \, \mathrm{d}\theta \, . \tag{6}$$

Note that the product term represents the joint predictive distribution in the case that parameter θ is known. The Bayesian predictive distribution simply averages (integrates) this with respect to the posterior probability distribution for θ. Hence, uncertainty about θ in light of the observed data is fully integrated.

In a Bayesian analysis we obtain forecasts and associated measures of uncertainty by calculating marginal probability distributions for quantities of interest by integrating the posterior distribution in (5) or the predictive distribution in (6). Performing these integrations analytically is typically not possible for realistically complex models such as those described above. Historically, this has prevented demographers and others from taking advantages of Bayesian methods for statistical inference. Recent developments in Bayesian computation have focussed on Markov chain Monte Carlo (MCMC) generation of samples from distributions such as (5) or (6); see Gelman et al. (2004) for details. Once a sample has been obtained from a joint distribution, then a sample from a distribution of any component or function of components is readily available. To generate samples from the posterior and predictive distribution in this paper, we used an MCMC sampling approach implemented using the WinBUGS software (Lunn et.al. 2000).

3.3 Model Uncertainty

In practical population forecasting, it is unrealistic for the analyst to be sure that any particular statistical model is the right one upon which to base their forecasts. Hence, the statistical methodology adapted should be one which allows for model uncertainty. Furthermore, we consider it essential that the measures of uncertainty associated with any forecast should incorporate both the uncertainty concerning the model and the uncertainty concerning the parameters of each model. In this paper, model uncertainty is directly integrated with parameter uncertainty into a single predictive probability distribution.

Formally, let $m = 1,\ldots,M$ index the models under consideration and let θ_m represent the parameter associated with model *m*. Note that different models may have parameters of different dimensionality. For example, the AR(1) model with non-zero mean has a three-dimensional parameter (μ, ϕ_1, σ^2). The likelihood function for model *m* is $f(y_{\{T\}} \mid \theta_m, m)$, the prior distribution for θ_m is $f(\theta_m \mid m)$ and the posterior distribution is

$$f(\theta_m \mid y_{\{T\}}, m) = \frac{f(y_{\{T\}} \mid \theta_m, m) f(\theta_m, m)}{f(y_{\{T\}} \mid m)}, \qquad (7)$$

where $f(y_{\{T\}} \mid m)$ is a normalising constant, known as the marginal likelihood for model *m* and is given by

$$f(y_{\{T\}} \mid m) = \int f(\theta_m \mid m) f(y_{T+K} \mid \theta_m, m) \mathrm{d}\theta . \qquad (8)$$

Prior uncertainty about models is encapsulated by a discrete probability distribution, $f(m)$, $m = 1,\ldots,M$. The prior model probabilities are all usually given the same values, $1/M$.

The posterior probability distribution for *m* given observed data $y_{\{T\}}$ is obtained by using Bayes Theorem as follows:

$$f(m \mid y_{\{T\}}) = \frac{f(y_{\{T\}} \mid m) f(m)}{f(y_{\{T\}})}. \qquad (9)$$

Hence, the posterior model probability for any model *m* is proportional to the product of the prior model probability and the marginal likelihood. Therefore, efficient computation of marginal likelihoods is essential for Bayesian inference under model uncertainty. See, for example, those described in O'Hagan and Forster (2004). In our implementation, we use a bridge sampler (Meng and Wong, 1996).

Finally, to obtain a predictive distribution for population forecasts in the presence of model uncertainty, (6) is extended to

$$\begin{aligned} f(y_{T+1},\ldots,y_{T+K} \mid y_{\{T\}}) &= \sum_{m=1}^{M} f(m \mid y_{\{T\}}) f(y_{T+1},\ldots,y_{T+K} \mid y_{\{T\}}, m) \\ &= \sum_{m=1}^{M} f(m \mid y_{\{T\}}) \int f(\theta_m \mid y_{\{T\}}, m) \prod_{k=1}^{K} f(y_{T+K} \mid y_{T+K-1}, \theta_m, m) \, \mathrm{d}\theta_m , \end{aligned}$$

which is the average of predictive distributions for individual models weighted by their posterior probabilities, $f(m \mid y_{\{T\}})$.

4. FORECASTS

In this section, we present parameter estimates from a range of individual AR models. In addition, the predictive probability distributions from these models are provided in order to gain a better understanding of the effect of expanding the dimensions of θ on future population growth rates. These individual forecasts are compared in the second subsection with a single forecast that accounts for our uncertainty in model choice. Finally, we compare our model averaged forecast against those produced by other means, namely classical time series methods and cohort component methods used by ONS.

4.1 Individual Models

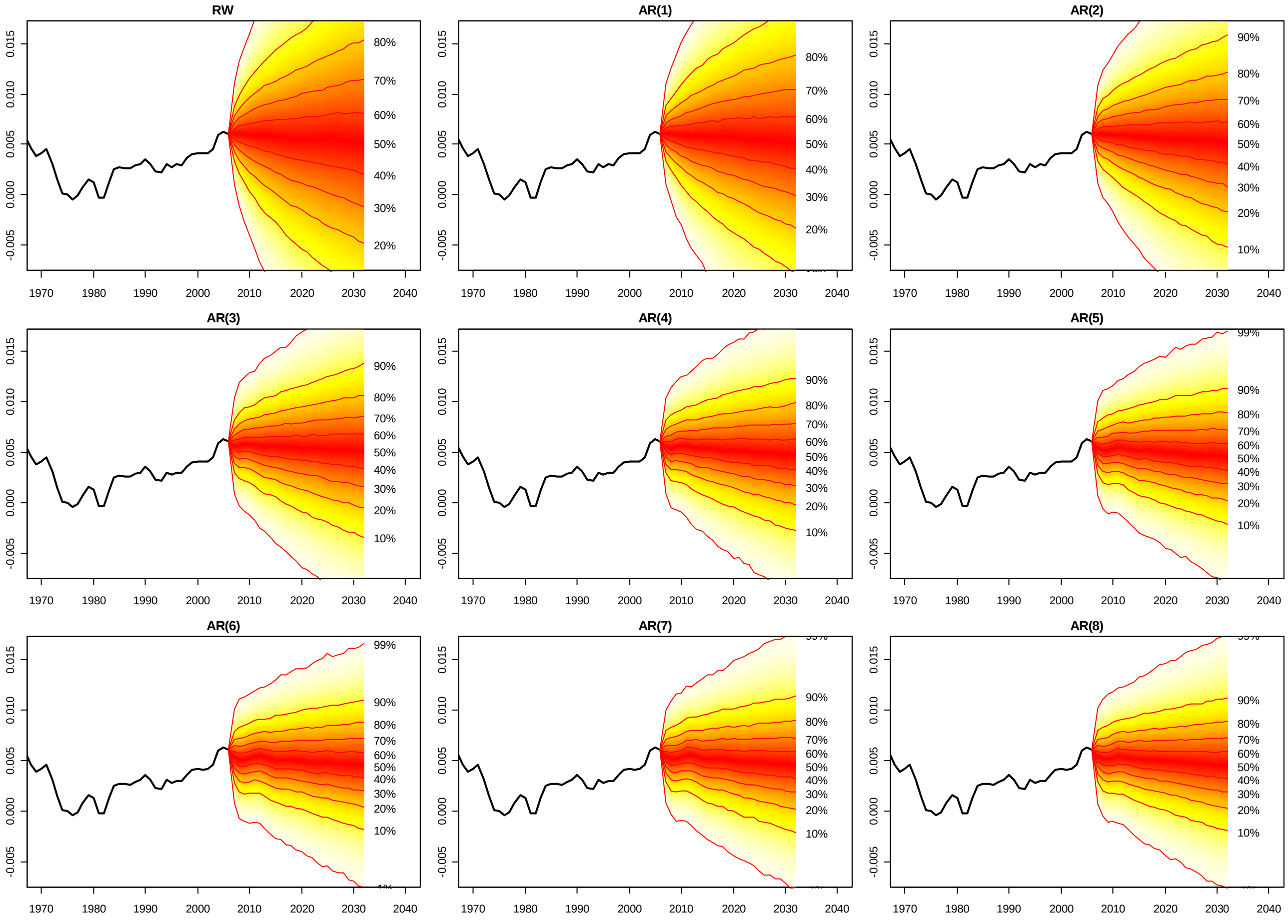
A set of nine models were considered for the differenced population growth rate, y_t, introduced in (2) and presented in the bottom panel of Figure 1. These consist of an independent normal (IN) model and eight autoregression models (with non-zero means), increasing in order from AR(1) to AR(8). This range of models was selected in order to represent all possible representations of autoregressive processes that might adequately describe the differences in the overall growth rate series. As we have no previous knowledge about the nature of the parameters in each model we assigned non-informative prior distributions: μ ~ N(0, 100), ϕ_k ~ N(0, 1) and σ ~ Uniform(0, 100). An MCMC sample of 10,000 observations was obtained from the posterior distribution for each model.

In a Bayesian analysis, marginal posterior distributions completely describe the uncertainty about each model parameter given the observed data. These are typically summarised using posterior means (as parameter estimates) and posterior standard deviations (as measures of uncertainty). The posterior means and standard deviations for the parameters of each of the nine models are presented in Table 1. Estimates of μ tend to be centred on zero with much lower standard deviations than their prior distributions. This feature was also true for the estimates of σ. In all models, the posterior means of ϕ_k at lower values of k were below zero, indicating negative autocorrelation for their respective lags. Estimates of ϕ_k , where $k > 5$ tend to be close to zero, signifying that the association between y_t and y_{t-k} becomes weak at larger values of k.

Table 1: Posterior Means and Standard Deviations of Model Parameters from MCMC Simulations and Model Comparisons Statistics.

m	μ	σ	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	ϕ_6	ϕ_7	ϕ_8	AIC	BIC	$f(m\mid y_{\{T}}$
IN	-0.00003 (0.00017)	0.00215 (0.00012)									-1558.75	-1556.65	0.03515
AR (1)	-0.00003 (0.00017)	0.00213 (0.00012)	-0.1536 (0.07858)								-1561.05	-1554.84	0.01803
AR (2)	-0.00004 (0.00016)	0.00208 (0.00011)	-0.1913 (0.07848)	-0.23275 (0.07776)							-1568.75	-1558.43	0.12161
AR (3)	-0.00006 (0.00016)	0.00203 (0.00011)	-0.24421 (0.07692)	-0.27889 (0.07668)	-0.22221 (0.07801)						-1575.29	-1560.87	0.54100
AR (4)	-0.00007 (0.00016)	0.00202 (0.00011)	-0.27600 (0.08036)	-0.31548 (0.07917)	-0.25796 (0.08042)	-0.1323 (0.07963)					-1576.61	-1558.08	0.17431
AR (5)	-0.00008 (0.00016)	0.00200 (0.00011)	-0.29726 (0.08085)	-0.35671 (0.08206)	-0.30694 (0.08257)	-0.17762 (0.08225)	-0.15658 (0.08042)				-1578.81	-1566.18	0.09724
AR (6)	-0.00009 (0.00016)	0.00200 (0.00011)	-0.30782 (0.08120)	-0.36911 (0.08393)	-0.32842 (0.08695)	-0.20312 (0.08828)	-0.17720 (0.08274)	-0.06892 (0.08115)			-1577.63	-1550.89	0.01131
AR (7)	-0.00009 (0.00016)	0.00200 (0.00011)	-0.30333 (0.08168)	-0.35741 (0.08478)	-0.31624 (0.08925)	-0.18207 (0.09171)	-0.15403 (0.09028)	-0.04924 (0.08455)	0.06089 (0.08002)		-1576.20	-1545.35	0.00121
AR (8)	-0.00009 (0.00016)	0.00200 (0.00011)	-0.29960 (0.08046)	-0.36279 (0.08516)	-0.32525 (0.08934)	-0.19299 (0.09333)	-0.17334 (0.09382)	-0.07220 (0.09117)	0.04309 (0.08579)	-0.06296 (0.08195)	-1574.87	-1543992	0.00013

Figure 2: Posterior Predictive Plots of Population Growth Rates


RW
AR(1)
AR(2)
AR(3)
AR(4)
AR(5)
AR(6)
AR(7)
AR(8)

Posterior predictive plots of the forecasted r_t from all nine models are illustrated in Figure 2. These are obtained from the forecast of y_t by rearranging (2) for each set of iterates and assuming $r_{2006} = 0.00609$ as in the data. Each shade of the forecasted fan represents a single percentile of the estimated posterior density, where darkest shades correspond to most central values and the lighter shades to the tails of the distribution. Contour lines are also plotted at each decile and the 1st and 99th percentile. Forecasts from the simple independent normal model provide a far greater level of uncertainty of future values. As the order of the AR models is expanded the posterior predictive distribution become comparatively tighter. As noted previously, ϕ_k for $k > 5$ are close to zero in the higher order AR models. This results in similar posterior predictive distribution for the higher order models, whereby the increase in the number of lagged terms no longer substantially reduces the width of the predictive distribution.

4.2 Weighted Model

The bridge sampler was used to calculate normalising constants for each model, from which the posterior model probabilities $f(m \mid y_{\{T\}})$ in Table 1 can be easily derived. These results indicate that over half of the probability rests on the AR(3) model. The next most likely model is the AR(4) model, followed by the AR(2) and AR(5) models. Higher ordered models appear very unlikely, as do the most basic models, with model probabilities below 0.05.

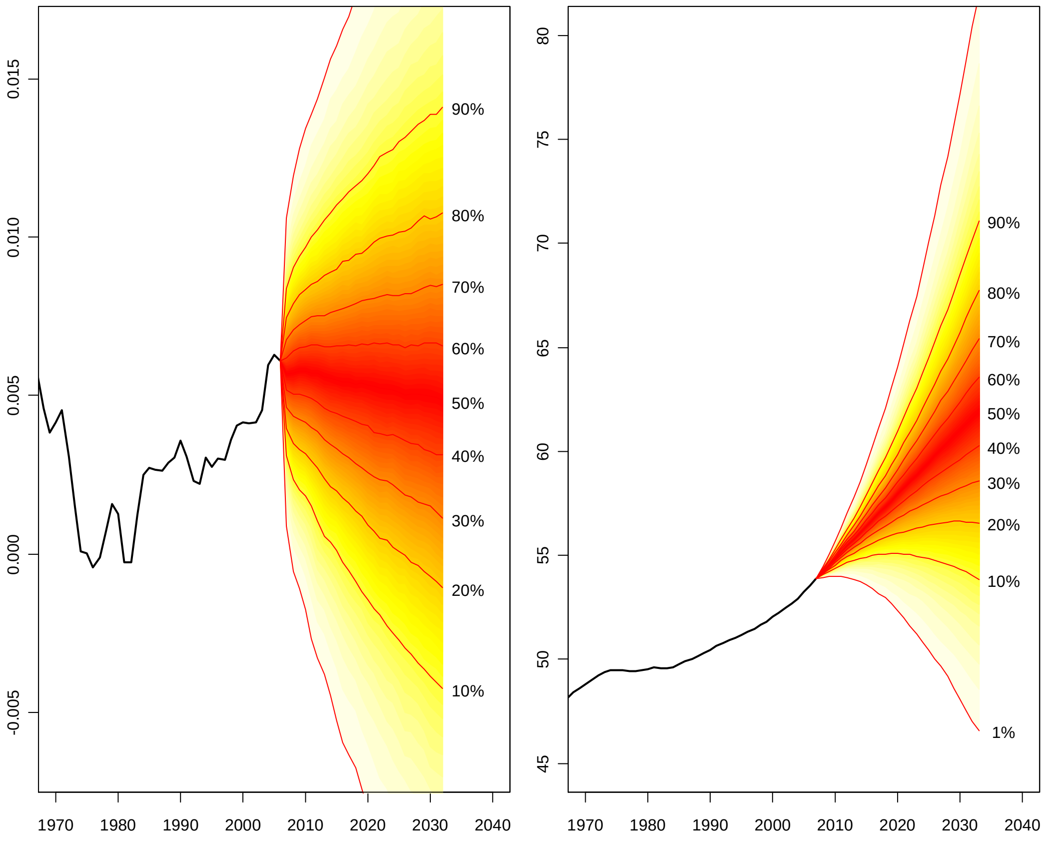
The predicted probability distribution of r_t averaged over all models, given the model probabilities, are presented in the left hand panel of Figure 3. Because a sample from the posterior of predicted probability distribution of each individual model is generated in the analysis, calculation of the averaged predicted probability distribution is trivial. Unsurprisingly this plot bares a large resemblance to the forecasts in Figure 2 which had large posterior model probabilities. On the right hand panel of Figure 3 we also present the resulting population forecast from the predicted probability distribution of r_t. Our results provide a median predicted population of 61.9 million in 2033. Numerous measures of uncertainty are also available, for example in 2033 the 20th percentile is 56.6 million and the 80th percentile is 67.7 million.

4.3 Comparison with Alternate Methods

For comparative purposes, traditional frequentist time series models corresponding to the nine models fitted above were estimated using the arima function in R 2.10.1 (R Development Core Team, 2010). Estimates of ϕ_k in all models were within 0.1 of the mean values presented in Table 1. Estimates from the arima function of μ and σ were also very similar (to a higher degree of accuracy) to those estimated using the Bayesian methodology. The close correspondences between parameter estimates are due to the reliance on data, rather than the (uninformative) priors, in the calculation of posterior distributions.

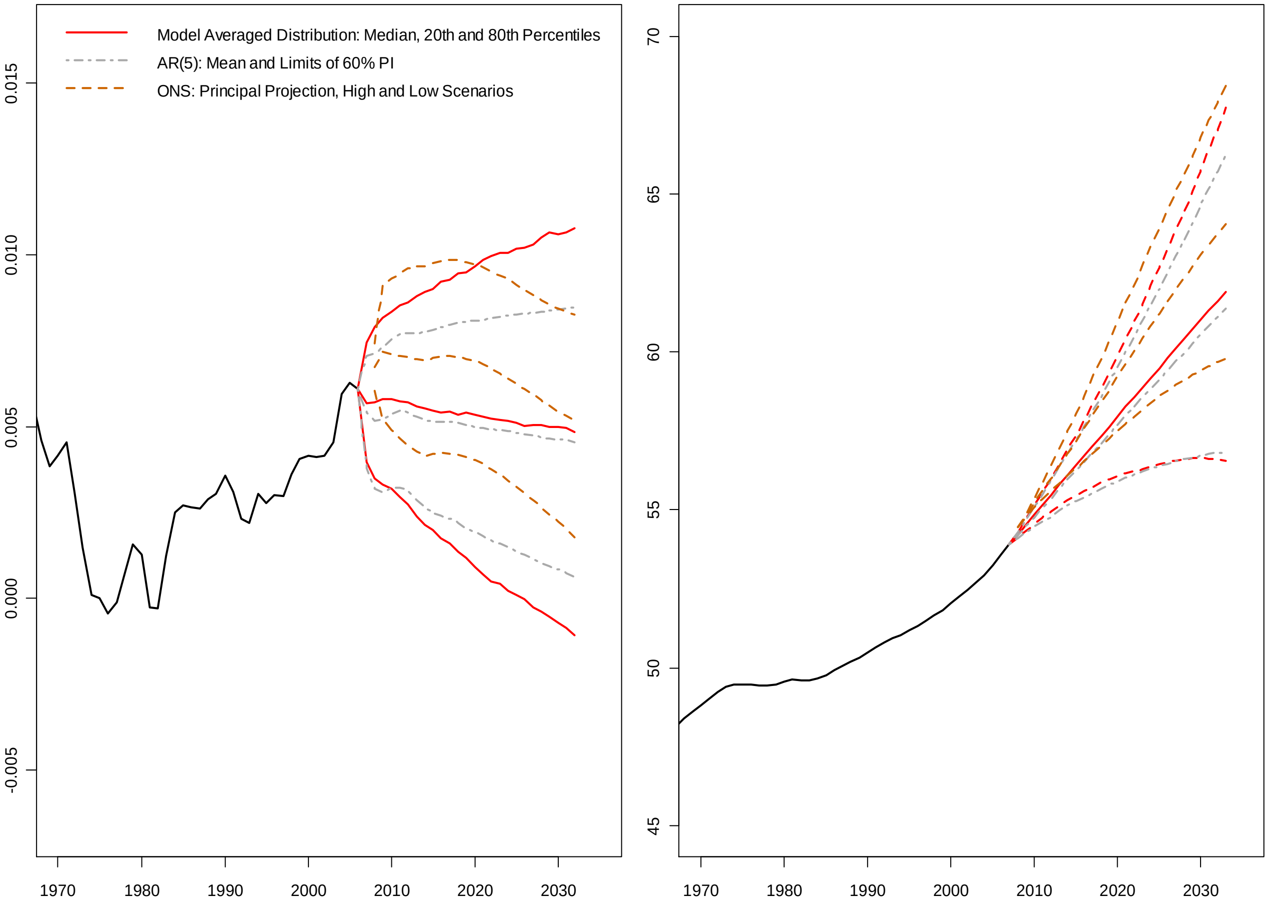
Model summary statistics from the models fitted in R are also provided in Table 1. The Akaike Information Criteria (AIC) of Akaike (1973) is commonly used for model selection for time series methods (Chatfield p256, 2004). This criteria favoured the AR(5) model, as opposed the model probabilities calculated using the bridge sampler which provided this model with a probability of 0.097. Hence, if we were to use the AIC as an alternative method for model selection in a frequentist setting, only a single model, with an estimated low probability, would be selected. The Bayesian Information Criteria (BIC) of Schwarz (1978), which penalises the inclusion of extra parameters more severely, is also presented in Table 1. This criteria closely resembles the posterior model probabilities and suggests AR(3) as a suitable model.

Figure 3: Joint Predicted Probability Distribution of the Model Averaged Growth Rates (left) and Resulting Population Forecast in Millions (right).



In Figure 4 we compare the results of the choice of a single model, based on the AIC, against our model averaged forecast. In left hand panel the mean forecast of r_t from the AR(5) model is displayed using the dot-dashed line. This was calculated using the predict.arima function in R. In addition, the 60% prediction intervals are also plotted alongside the 20th, 50th and 80th percentiles of model averaged density (solid lines). Comparisons of the two measures of central tendency are very similar. However, the AR(5) model provides a smaller amount of uncertainty, from both the parameter estimation and model selection process, when using traditional frequentist time series methods. These similarities and differences are also partly reflected in the forecasts of the total population plotted on the right hand panel of Figure 4. The mean forecasted population for the AR(5) model is 61.4 million in 2033, whilst the upper and lower limits for the 60% prediction intervals are 56.8 and 66.3 million. The closeness in the lower prediction interval with the 20th percentile is due to the lower forecasted values of r_t, including a larger fall in the earlier years of the forecast. This early drop results in subsequent values of lower prediction interval of p_t to remain below the 20th percentile until 2028.

Figure 4: Comparison of Alternative Future Growth Rates (left) and Population in Millions (right).



In the United Kingdom, the ONS prepare a set of projected total populations estimated using a cohort component methodology under a range of deterministic scenarios. In this paper we focus on three variants (principal, high and low) published in the latest set of projections for England and Wales (ONS, 2009). All three variants are based on the sets of demographic trend-based assumptions for future fertility, mortality and net migration. The principal variant relies on assumptions considered to best reflect demographic patterns at the time they were adopted. The high (or low) population variant assumes a combination of high (or low) fertility, life expectancy and net migration. They are intended to provide users with a better understanding of future uncertainty in population change. All three variants of population totals are displayed on the right hand panel in Figure 4. On the left hand panel are the derived values of r_t. The central, dashed line represents the principal projections, whilst the upper and lower dashed line represent the high and low population variants respectively. The panels in Figure 4 illustrate a number of differences between the ONS principal projection and that of our model averaged forecasts. First, the uncertainty in the ONS rate, represented by their high and low variants, is far smaller than that of our model averaged forecasts at all points of time. Second, the uncertainty in the rate of population growth of the ONS projection does not increase substantially over time. Third, the ONS principal population projection in 2033 of 63.7 million is slightly higher than our model averaged median (61.9 million), despite a reduction in the rate toward the median of the model averaged forecast towards the end of the horizon. This feature is caused by an assumption of higher growth rates throughout the future period. Finally, the high and low variants in the projected population totals by the ONS lie within the 81st and 36th percentiles of the posterior predictive distribution.

5. CONCLUSION

In this paper we have demonstrated the use of Bayesian time series methods for the forecasting of the future population of England and Wales using a historical series of population growth rates. The forecasts have explicitly allowed for uncertainties in the data, parameters of the model and the model itself by using probability distributions, which are fully represented in the final probabilistic population forecast.

All of the simple Bayesian time series models assumed stationarity in the y_t. However, the bottom panel of Figure 3 indicates that there is some degree of volatility in the differenced population growth rates. More complex time series models exist, such as stochastic volatility models that allow the variance of z_t to be time-dependent. Such models replace σ^2 in (4) with σ_t^2, where a time series model, typically a AR(1) process, is then specified for log σ_t^2. Accounting for this heterogeneity will allow for forecasts to adjust to the level of volatility estimated in the jump off period. Further extensions to the modelling of the growth rate can also be explored by decomposing r_t to demographic components of population change. Separate series of births, deaths and migration can be modelled as a multivariate process using Bayesian Vector Autoregressive (VAR) models. We are currently investigating both of these extensions.

Simple time series models were used in this paper to forecast future population growth. The median of our predictive distribution for future populations are slightly lower, but not drastically different to, the principal projection estimated by ONS using a more complex cohort component methodology. Such methodologies require a large amount of data on current age and sex structure and numerous assumptions on rates of demographic components. However, unlike the more complex cohort component method the forecasting methods used in this paper are able to quantify our uncertainty through a posterior predictive distribution.

Our model averaged posterior predictive distribution tended to be wider than those provided by prediction intervals from traditional frequentist time series methods. This is not unexpected as intervals for a single model selected on the basis of a model fit statistics (such as the AIC or BIC) will tend to be too narrow (Chatfield p86, 2004). Causes of these differences are include; uncertainty about the model and a changing environment. Thus, the use of model averaging allows a more realistic picture of the uncertainty of future population to be obtained. In this paper we used the bridge sampler to calculate the normalising constant for each model, and then derive model probabilities. This method can also be applied to deal with a wider range of models including the extensions previously mentioned in this section.

As Booth (2006) notes; the incorporation of informed judgements have formed the basis of many of the assumptions in traditional population projections, either as inputs or in combination with extrapolation or as a sole input. However methods tend to be unsystematic or inadequately documented, even in developed countries. The Bayesian approach allows uncertainty in the data, model parameters and model selection to be fully quantified using probability distributions.

In summary, the use of time series methods for population growth rates offer a simple alternative method to forecast population. In conjunction with Bayesian inference, population forecasts are able to account for multiple sources of uncertainty including data, parameter estimates and model selection. Consequently population forecasts may contain more realistic measurements of uncertainty and thus improve users' potential for planning and understanding population change.

6. REFERENCES

Ahlburg, Dennis. 1987. Population Forecasts for South Pacific Nations using Autoregressive Models 1985 - 2000. *Journal of Population Research*, 4(2), 157–167.

Akaike, H. 1973. Information Theory and an Extension of the Maximum Likelihood Principle. *Pages 267–281 of: International Symposium on Information Theory, 2nd, Tsahkadsor, Armenian SSR*.

Alho, Juha M., & Spencer, Bruce D. 1985. Uncertain Population Forecasting. *Journal of the American Statistical Association*, 80(390), 306–314.

Alho, Juha M., & Spencer, Bruce D. 2005. *Statistical Demography and Forecasting (Springer Series in Statistics)*. Springer.

Booth, H. 2006. Demographic forecasting: 1980 to 2005 in review. *International Journal of Forecasting*, 22(3), 547–581.

Chatfield, Chris. 2003. *The Analysis of Time Series: An Introduction, Sixth Edition (Chapman & Hall/CRC Texts in Statistical Science)*. 6 edn. Chapman and Hall/CRC.

Coleman, David, & Salt, John. 1992. *The British Population: Patterns, Trends, and Processes*. Oxford University Press, USA.

Gelman, Andrew, Carlin, John B., Stern, Hal S., & Rubin, Donald B. 2003. *Bayesian Data Analysis, Second Edition (Texts in Statistical Science)*. 2 edn. Boca Raton, USA: Chapman & Hall/CRC.

Hinde, Andrew. 2003. *England's Population: A History since the Domesday Survey (Arnold Publication)*. Bloomsbury USA.

Human Mortality Database, 2009. *University of California, Berkeley, USA, and Max Planck Institute for Demographic Research in Rostock, German.* Retrieved 15th September 2009 URL: http://www.mortality.org/

Keyfitz, Nathan. 1991. Uncertainty in National Population Forecasting: Issues, Backgrounds, Analyses, Recommendations. *Population Studies: A Journal of Demography*, 45(3), 545–546.

Lee, Ronald D. 1998. Probabilistic Approaches to Population Forecasting. *Population and Development Review*, 24, 156–190.

Lunn, David J., Thomas, Andrew, Best, Nicky, & Spiegelhalter, David. 2000. WinBUGS – A Bayesian modelling framework: Concepts, structure, and extensibility. *Statistics and Computing*, 10(4), 325–337.

Lynch, Scott M. 2007. *Introduction to Applied Bayesian Statistics and Estimation for Social Scientists (Statistics for Social and Behavioral Sciences)*. 1 edn. Springer.

McNown, Robert, Rogers, Andrei, & Little, Jani. 1995. Simplicity and complexity in extrapolative population forecasting models. *Mathematical Population Studies: An International Journal of Mathematical Demography*, 5(3), 235–257.

Meng, Xiao-Li, & Wong, Wing H. 1996. Simulating Ratios of Normalizing Constants via a Simple Identity: A Theoretical Exploration. *Statistica Sinica*, 6, 831–860.

Office for National Statistics, 2009. *Statistical Bulletin: National Population Projections, 2008-base*. Retrieved 30th October 2009. URL: http://www.statistics.gov.uk/pdfdir/pproj1009.pdf

O'Hagan, Antony, & Forster, Jon. 2004. *The Advanced Theory of Statistics, Vol. 2B:Bayesian Inference*. 2 edn. Wiley.

Pflaumer, P. 1992. Forecasting US population totals with the Box-Jenkins approach. *International Journal of Forecasting*, 8(3), 329–338.

R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Saboia, Joao L. M. 1974. Modeling and Forecasting Populations by Time Series: The Swedish Case. *Demography*, 11(3), 483–492.

Schwarz, Gideon. 1978. Estimating the Dimension of a Model. *The Annals of Statistics*, 6(2), 461–464.